The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2017-268-RC3, 2018

Interactive comment on "Glacio-hydrological melt and runoff modelling: a limits of acceptability framework for model selection" *by* Jonathan D. Mackay et al.

Anonymous Referee #3

Received and published: 6 March 2018

Summary of the manuscript This manuscript (ms) aims to presents "A limits of acceptability framework for model selection for glacio-hydrological melt and runoff modelling". For this purpose three model structures were calibrated and validated with ice volume change observations, satellite derived snow cover images and runoff data. The authors conclude that "it remains to be seen if the framework can be used", acknowledging that the results were "not necessarily more consistent across the full range of signatures". Further studies would be needed to investigate these conclusions.

Evaluation I think the topic of this manuscript is highly relevant and important for hydrological modelling and therefore it should be tackled thoroughly and with care. In recent

years many studies have compared model complexity and different observational data sets to investigate glacio-hydrological melt and runoff modelling. Accordingly, I would highly appreciate a concise and well-established framework of acceptability of glacio-hydrological models. However, I have my major doubts about the present ms and if it delivers what the title promises. My concerns are the following: i) the authors conclude that the results were "not necessarily more consistent across the full range of signatures" and that "it remains to be seen if the framework can be used". Based on the presented results I would agree with these conclusions and will provide comments how this can be tacked in a more thorough way below. But I also would argue that a framework that cannot provide consistent results and has to be investigated further is not publishable. ii) The evaluation metrics (equation 9) are based on user defined signatures thresholds, rather than well-established efficiency criteria. This makes it impossible for the reader to assess the performance of the model. iii) Inter-annual average altitude dependent simulated and observed snow cover is presented (Fig 12) rather than daily time steps. This makes it impossible to assess how the model performs during years with enhanced snowfall and reduced snowfall. Averages of inter-annual performances dilute the performance, making it impossible for the reader to assess the real performance. iv) Same argument is valid for runoff: in my opinion daily time-steps should be compared for a calibration and validation period. v) In my opinion, a validation of the modelling results based on the presented "user defined" evaluation metrics is inadequate, because it is not objective (If other thresholds were selected the results might have been completely different, making the presented analysis subjective). vi) Many of the concerns addressed above have been analyzed and discussed in recent works, some of which are referenced in this ms but not taken adequately into account. In my opinion the authors fail to connect to previous works and point out why this framework is novel and how this ms fills an existing knowledge gap.

I leave it up to the editors of "The Cryosphere" to decide if the present ms can be revised or should be resubmitted as a new ms. I would recommend to address the following concerns prior to publication:

Major concerns: 1) A framework has to work to be published. Accordingly, the framework should be able to reproduce adequately i) bi-annual glacier mass balances (accumulation and ablation phase) over several years, ii) daily snow cover ratio over several years to demonstrate that it works for snow intense and snow poor years. iii) daily runoff patterns. 2) If three model structures are used and they produce the same results, I would argue that the model structure is insensitive to the modelling approach. Nevertheless, the title claims to provide an acceptability framework for model structure. How can the framework say anything about model structure, it all the structures produce similar results? 3) Evaluation metrics should be consistent and comparable with previous literature. I understand that the chosen study site might lead to exceptionally low performance due to extreme weather patterns, but I would still use Nash-values for runoff, RMSE for glacier mass balances and a representative index for the snow cover area. This would enable a comparison of the model performance with previous works. 4) In my opinion, the ms can be significantly shortened, made more concise and the literature should be selected more carefully.

Specific recommendations: âĂć Title: I find the title confusing and misleading. Please provide a more focused title. âĂć Abstract: do models underpin the understanding of future climate change? I would argue that only the analysis of the results obtained with models can do that (it's a detail, but important, I recommend that the entire ms is revised and such flaw statements are corrected) âĂć Introduction: in my opinion this chapter can be significantly shortened to a third of its current length. But I also think that it should be more targeted and the cited literature should be more focused. âĂć Referencing: in my opinion 3 references are sufficient for one statement; in this ms up to 17 references are used of a single statement. This is misleading and not helpful for the reader. I think every references should be carefully chosen and only the most relevant references should be used (this would keep the reader focused on the topic of the ms). Also, I would recommend referencing the first publication that has investigated a topic, rather than referencing newer articles who have just build on previous works. Referencing is a tedious job, involves a lot of reading but should be taken seri-

ously. âĂć I would tend to disagree that this is the "first kind to apply a signature-based limits of acceptability (LOA) framework". In my opinion numerous studies have done this, simply in a different manner. âĂć Figure 1: Why are meteo-data presented as an inlet in the map of the study site? Please stick to a coherent structure and provide all figures in a standard scientific style. âĂć Observational data: in my opinion this chapter can also be significantly shortened. None of the data were collected by the authors, so the methods how they were collected can be found in the relevant literature. âĂć Glacio-hydrological model: has the model never been published before? Does it have a name? Why not use a model that is well established? Then the description of the model could be shortened and the focus could be on the framework. âĂć Driving climate data: The Icelandic Met Office (IMO) provides gridded reanalysis data (P, T and I). They are continuously updating their gridded data sets. Accordingly, I would recommend using the official data set of the IMO, rather than reproducing something that has been investigated for many years by the IMO. âĂć Figure 3: It is impossible to assess how well the correction worked based on the presented data; see my previous comment. I recommend computing some relevant statistical values (see some of the cited literature). âĂć Fig 4,5,67: please see my concern regarding aggregated monthly or inter-annual comparison of observed and simulated runoff and snow cover; What is the use of modelling daily time steps if only mean monthly results are evaluated? If the objective is only the get monthly means correctly than I recommend using a monthly time step in the modelling framework. âĂć Fig 8: see my comment regarding the evaluation metrics; âĂć Fig 9: I recommend making a table rather than a figure; âĂć Fig 10: why select a case study which has only two glacier volume observations and 23 year of data gap between the two observations? There are numerous case studies which have bi-annual glacier mass balances, which would be a lot more valuable to establish a framework of acceptability for glacio-hydrological models; If the authors want to establish a transparent framework, I recommend selecting a case study with enough observational data to test the framework thoroughly. âĂć Fig 11: in my opinion these results suggest that the framework does not work; none of the simulations are acceptable, even according to the standards set by the authors; âĂć Fig 12: this figure illustrated the discrepancy between observed and simulated snow cover: In my opinion a acceptability framework should be able to identify if a model works during snow intensive and snow poor years. Here we only see that it never really works.

Final remark: I do think that this study can become an important contribution to hydrological modeling. However, all of the comments above would need to be accounted for and/or addressed. I would like to encourage the authors to have a thorough discussion on the purpose of this study. Furthermore, I encourage the authors to read some of the literature that is already referenced; many of the concerns addressed above have already been investigated and solutions have been presented. I would like to wish the authors lots of success and good luck with further research on such a framework.

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2017-268, 2018.